A priori estimations of a global homotopy residue continuation method.

I. Gruais[1], N.M.M. Cousin–Rittemard[1*], H.A. Dijkstra[2]

[1]Institut de Recherche Mathématique de Rennes,Université de Rennes 1, Rennes, France.
[2]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, 80523-1371, United States of America.

Abstract

This work is concerned with the a priori estimations of a global homotopy residue continuation method starting from a disjoint initial guess. Explicit conditions ensuring the quadratic convergence of the underlying Newton-Raphson algorithm are proved.

Key words: continuation methods; error estimate; global homotopy; residue.

1 Introduction

The models of nonlinear physical phenomena depend on parameters says as $\mu \in \mathbb{R}^p$. The Dynamical Systems Theory studies the features of the transitions in these nonlinear systems. This theory is basically comprised of the bifurcation theory and the theory of ergodic systems. One of the important basic issues of the bifurcation theory is the determination of the fixed points of the system under investigation, says as

$$\frac{\partial \mathbf{u}}{\partial t} \equiv \mathbf{A}(\mathbf{u}, \mu) = \mathbf{0} \tag{1}$$

Here, $\mathbf{u} \in \mathbb{R}^n$ is the vector of the unknowns and $\mathbf{A} : \mathbb{R}^{\mathbf{n}} \times \mathbb{R}^{\mathbf{p}} \to \mathbb{R}^{\mathbf{n}}$ is a nonlinear operator.

The branches of steady states are computed versus a control parameter $\mu \in \{\mu_1, \mu_2, ..., \mu_p\}$ using the continuation methods [7]. A vast and rich literature exists: see [1, 2, 3, 4, 5, 9, 11] for example.

As it is well known, the Newton–Raphson and the pseudo-arclength continuation methods can be described as below. The Newton–Raphson method

*Email: nathalie.rittemard@univ-rennes.fr

consists in computing a branch of fixed points of (1) through infinitesimal increments of a control parameter $\mu \in \{\mu_1, \mu_2, ..., \mu_p\}$, says as

$$\mathbf{D_u A}^k \; \delta\mathbf{u^k} = -\mathbf{A^k} \quad , \quad \mathbf{u^k} \leftarrow \mathbf{u^k} + \delta\mathbf{u^k} \tag{2}$$

where $\delta\mu$ is a small increment, $\delta\mathbf{u^k} = \mathbf{u^{k+1}} - \mathbf{u^k}$, $\mathbf{D_u A}^k$ is the Jacobian matrix of $\mathbf{A}^k$ with respect to $\mathbf{u}$ at the kth estimate $(\mathbf{u}^k, \mu + \delta\mu)$. At regular points, it holds that

$$\operatorname{rank}\left(\mathbf{D_u A}^k\right) = n. \tag{3}$$

The pseudo–arclength continuation method consists in calculating a parametrized branch of fixed points through infinitesimal increments of the curvilinear abscissa s, says as

$$\begin{pmatrix} \mathbf{D_u A}^k & \mathbf{D}_\mu \mathbf{A}^k \\ \mathbf{D_u N}^{kT} & \mathbf{D}_\mu \mathbf{N}^k \end{pmatrix} \begin{pmatrix} \delta\mathbf{u^k} \\ \delta\mu^k \end{pmatrix} = - \begin{pmatrix} \mathbf{A}^k \\ \mathcal{N}^k \end{pmatrix} \;, \tag{4a}$$

$$\begin{pmatrix} \mathbf{u}^k \\ \mu^\mathbf{k} \end{pmatrix} \leftarrow \begin{pmatrix} \mathbf{u}^k + \delta\mathbf{u^k} \\ \mu^\mathbf{k} + \delta\mu^k \end{pmatrix} \tag{4b}$$

with the scalar normalization

$$\mathcal{N}\left(\mathbf{u}(s), \mu(s)\right) \equiv \mathbf{D_u N}.\delta\mathbf{u} + \mathbf{D}_\mu\mathbf{N}\;\delta\mu - \delta\mathbf{s} = \mathbf{0} \tag{5}$$

where $\mathbf{D}_\mu\mathbf{N}$ and $\mathbf{D_u N}$ are derivatives of the operator $\mathcal{N}$ with respect to μ and $\mathbf{u}$ respectively and $\mathbf{D}_\mu\mathbf{A}^k$, $\mathbf{D}_\mu\mathbf{N}^k$, $\mathbf{D_u N}^k$ are the derivatives of the operators at the current estimate $\left(\mathbf{u}^k, \mu^k\right)$.

In this contribution, a priori estimations of a global homotopy continuation of the residue are presented. Explicit conditions ensuring the quadratic convergence of the Newton-Raphson and the pseudo–arclength continuation methods are derived.

The residue continuation method is based on the global homotopy concept which was pioneered in [8]. Let $(\mathbf{u}^*, \mu^*)$ be an initial guess of a disjoint fixed point $(\mathbf{u}_0, \mu_0)$ of (1). The idea of the residue continuation method is to solve the global homotopy :

$$\mathcal{H}\left(\mathbf{u}, \mu, \alpha\right) \equiv \mathbf{A}(\mathbf{u}, \mu) - \alpha\,\mathbf{r} = 0 \tag{6a}$$

$$\mathcal{K}\left(\mathbf{u}, \mu\right) \equiv k_\mathbf{u}.\mathbf{u} + k_\mu \mu = 0 \tag{6b}$$

where $\mathbf{r} = \mathbf{A}(\mathbf{u}^*, \mu^*)$ is the *residue* and α is the *residue parameter*.

For a given $\mathbf{r}$, and assuming that

$$k_\mu \neq 0, \tag{7}$$

it follows from the Implicit Function Theorem that (6) can be written as

$$\mathcal{H}\left(\mathbf{u}(\alpha)\right) \equiv \mathbf{A}(\mathbf{u}(\alpha), \mu(\mathbf{u}(\alpha))) - \alpha\,\mathbf{r} = 0 \tag{8a}$$

$$\mathcal{K}\left(\mathbf{u}(\alpha)\right) \equiv k_\mathbf{u}.\mathbf{u}(\alpha) + k_\mu\,\mu(\mathbf{u}(\alpha)) = 0 \tag{8b}$$

Let $(\alpha_\nu)_\nu$ be a real sequence such that $\alpha_\nu \in I \equiv [a, b] \subset \mathbb{R}$. For $(\mathbf{u}_{\nu-1}, \mu_{\nu-1})$ solution of the homotopy (6), we note

$$\mathbf{r}_{\nu-1} = \mathbf{A}(\mathbf{u}_{\nu-1}, \mu_{\nu-1}). \tag{9}$$

The Newton-Raphson method is used to solve the system of equations (8), says as

$$\mathbf{D_u}\mathcal{H}_\nu^k \,\delta\mathbf{u}_\nu^\mathbf{k} = -\mathcal{H}_\nu^\mathbf{k} \quad , \quad \mathbf{u}_\nu^\mathbf{k} \leftarrow \mathbf{u}_\nu^\mathbf{k} + \delta\mathbf{u}_\nu^\mathbf{k} \tag{10}$$

while for the pseudo-arclength continuation, it becomes

$$\begin{pmatrix} \mathbf{D_u}\mathcal{H}_\nu^k & D_\alpha\mathcal{H}_\nu^k \\ \mathbf{D_u}\mathbf{N}_\nu^{kT} & D_\alpha\mathbf{N}_\nu^k \end{pmatrix} \begin{pmatrix} \delta\mathbf{u}_\nu^k \\ \delta\alpha_\nu^k \end{pmatrix} = -\begin{pmatrix} \mathcal{H}_\nu^k \\ \mathcal{N}_\nu^k \end{pmatrix}, \tag{11a}$$

$$\begin{pmatrix} \mathbf{u}_\nu^k \\ \alpha_\nu^k \end{pmatrix} \leftarrow \begin{pmatrix} \mathbf{u}_\nu^k + \delta\mathbf{u}_\nu^k \\ \alpha_\nu^k + \delta\alpha_\nu^k \end{pmatrix} \tag{11b}$$

Here, μ_ν is such that $k_\mathbf{u}.\mathbf{u}_\nu + k_\mu\mu_\nu = 0$ and $\mathcal{H}_\nu^k \equiv \mathbf{A}_\nu^k - \alpha_\nu\mathbf{r}_{\nu-1}$. Furthermore, $\mathbf{D_u}\mathcal{H}_\nu^k$ and $D_\alpha\mathcal{H}_\nu^k$ are the Jacobians with respect to $\mathbf{u}$ and the residue parameter respectively, at the current estimate $(\mathbf{u}_\nu^k, \alpha_\nu^k)$ of the solution $(\mathbf{u}_\nu, \alpha_\nu)$ of (6).

From (8a) and (9), it follows that

$$|\alpha_\nu| = \frac{\|\mathbf{r}_\nu\|}{\|\mathbf{r}_{\nu-1}\|}. \tag{12}$$

Hence the residue parameter α_ν may be seen as the control parameter of the norm of the residue. The residue thus increases (decreases) as long as $|\alpha| > 1$ ($|\alpha| < 1$). It follows that $\alpha = 1$ is a critical value corresponding to an extremum of the norm of the residue.

2 A priori estimations

2.1 Newton–Raphson

For any subdivision $(\alpha_\nu)_\nu$ of $I \subset \mathbb{R}$, we denote by $\mathbf{u}_\nu$ the solution $\mathbf{u}(\alpha_\nu)$. Given an initial guess ($\mathbf{u}^0 \equiv \mathbf{u}_{\nu-1}$, $\mu_\nu^0 \equiv \mu_{\nu-1}$) solution of (6) with $\alpha \equiv \alpha_{\nu-1}$ and $\mathbf{r} \equiv \mathbf{r}_{\nu-2}$, the Newton-Raphson's scheme is written in the equivalent form:

$$\begin{aligned} &\text{For } \nu = 1, \cdots, N, \\ &\text{For } k = 0, \cdots, p_\nu - 1, \\ &\mathbf{D_u}\mathbf{A}_\nu^k(\mathbf{u}_\nu^{k+1} - \mathbf{u}_\nu^k) = -\mathbf{A}_\nu^k + \alpha_\nu\mathbf{r}_{\nu-1} \quad , \quad \mathbf{u}_\nu^k \leftarrow \mathbf{u}_\nu^{k+1} \end{aligned} \tag{13}$$

with the corresponding value of the control parameter μ_ν such that

$$k_\mathbf{u} \,.\, (\mathbf{u}_\nu - \mathbf{u}_{\nu-1}) + k_\mu(\mu_\nu - \mu_{\nu-1}) = 0. \tag{14}$$

Assuming that $\mathbf{D_u A}$ is nonsingular for every $\mathbf{u} \in \mathbb{R}^n$, the operator $\mathbf{A} : \mathbb{R}^n \mapsto \mathbb{R}^n$ (and hence also $\mathcal{H} : \mathbb{R}^n \mapsto \mathbb{R}^n$) is a homeomorphism. Therefore, for every α

in some compact range $I_\nu \subset I \subset \mathbb{R}$, with extremities α_ν and $\alpha_{\nu-1}$, the equation (6) admits a unique solution denoted by $\mathbf{u}(\alpha)$:

$$\mathbf{u}(\alpha) = \mathbf{A}^{-1}\left(\phi_\nu(\alpha)\mathbf{r}_{\nu-1}\right), \quad \phi_\nu(\alpha) \equiv \frac{(1-\alpha_\nu)\alpha - (1-\alpha_{\nu-1})\alpha_\nu}{\alpha_{\nu-1}-\alpha_\nu}\mathbf{r}_{\nu-1}. \tag{15}$$

As $\mathbf{A}^{-1}$ is continuously differentiable it follows that $\alpha \mapsto \mathbf{u}(\alpha)$ is continuous and piecewise C^1 while $\alpha \mapsto \mathbf{D_u A}(\mathbf{u}(\alpha))^{-1}$ is continuous on each I_ν. Therefore, assuming that the sequence $(\|\mathbf{r}_\nu\|)_\nu$ is bounded, there exists a constant $c > 0$ such that

$$\|\mathbf{D_u A}(\mathbf{u}(\alpha))^{-1}\| \leq c, \quad \forall \alpha \in I_\nu \subset I \subset \mathbb{R} \tag{16}$$

and it follows that

$$\mathbf{u}'(\alpha) = \mathbf{D_u A}(\mathbf{u}(\alpha))^{-1}\left(\frac{(1-\alpha_\nu)}{\alpha_{\nu-1}-\alpha_\nu}\mathbf{r}_{\nu-1}\right), \quad \alpha \in I_\nu. \tag{17}$$

We denote by $\mathcal{C}$ the limit curve of the Newton-Raphson process, defined as

$$\mathcal{C} \equiv \{\mathbf{u}(\alpha) \in \mathbb{R}^n, \quad \alpha \in I \subset \mathbb{R}\}. \tag{18}$$

As I is a compact convex set and $\mathbf{u}$ is piecewise C^1 and continuous, there exists a compact and convex set $D \subset \mathbb{R}^n$ such that $\mathcal{C} \subset D$.

Proposition 2.1 *For some constant $c' > 0$ we get:*

$$\forall \mathbf{u} \in D: \quad \|\mathbf{D_u A}(\mathbf{u})^{-1}\| \leq c'. \tag{19}$$

Proof. As $\mathbf{D_{uu}A}$ is continuous, it is also bounded on D and we define the constant $\kappa > 0$ as it follows:

$$\|\mathbf{D_{uu}A}\| \leq \kappa, \quad \forall \mathbf{u} \in D \tag{20}$$

Because $\mathbf{D_u A}$ is continuous on D, it is also uniformly continuous on D, and hence

$$\begin{gathered}\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall \mathbf{u}, \mathbf{u}' \in D, \\ \|\mathbf{u} - \mathbf{u}'\| < \eta \Longrightarrow \|\mathbf{D_u A}(\mathbf{u}) - \mathbf{D_u A}(\mathbf{u}')\| \leq \varepsilon.\end{gathered} \tag{21}$$

Furthermore, given any $\mathbf{u}, \mathbf{u}' \in D$ and setting

$$\mathbf{f(t)} \equiv \mathbf{D_u A}(\mathbf{u} + \mathbf{t}(\mathbf{u}' - \mathbf{u}))$$

we get, as f is continuously differentiable:

$$\mathbf{D_u A}(\mathbf{u}) - \mathbf{D_u A}(\mathbf{u}') = \mathbf{f(1)} - \mathbf{f(0)} = \int_{\mathbf{0}}^{\mathbf{1}} \mathbf{f'(t)dt}$$

$$= \int_0^1 \mathbf{D_{uu}A}(\mathbf{u} + t(\mathbf{u}' - \mathbf{u}))dt\ (\mathbf{u}' - \mathbf{u}).$$

Then, there holds

$$\forall \mathbf{u}, \mathbf{u}' \in D, \quad \|\mathbf{D_u A}(\mathbf{u}) - \mathbf{D_u A}(\mathbf{u}')\| =$$

$$\left\| \int_0^1 \mathbf{D_{uu}A}(\mathbf{u}+t(\mathbf{u}'-\mathbf{u}))dt\ (\mathbf{u}'-\mathbf{u}) \right\| \leq \kappa \|\mathbf{u}-\mathbf{u}'\|$$

so that we may choose $\eta = \dfrac{\varepsilon}{\kappa}$.

Moreover, for every $\mathbf{u}^0 \in D$, we find

$$\begin{aligned}
\|\mathbf{D_u A}(\mathbf{u})^{-1}\| &\leq \frac{\|\mathbf{D_u A}(\mathbf{u}^0)^{-1}\|}{1-\|\mathbf{D_u A}(\mathbf{u}^0)^{-1}\|\ \|\mathbf{D_u A}(\mathbf{u})-\mathbf{D_u A}(\mathbf{u}^0)\|} \\
&\leq \frac{\|\mathbf{D_u A}(\mathbf{u}^0)^{-1}\|}{1-\varepsilon\|\mathbf{D_u A}(\mathbf{u}^0)^{-1}\|}.
\end{aligned}$$

As a consequence, the equation (16) holds with $c' = \dfrac{c}{1-\varepsilon c}$ so that we choose

$$\varepsilon \in [0, \frac{1}{2c}]. \tag{22}$$

■

Following [6] (see [10]), with each sequence $(\mathbf{u}_\nu^k)_k$, we associate the quantities $\beta_{k,\nu}$, $\eta_{k,\nu}$, $\gamma_{k,\nu}$, $t_{k,\nu}^{\pm}$ according to the recurrence introduced as it follows. Let $\beta_{0,\nu}$, $\eta_{0,\nu}$, $\gamma_{0,\nu}$, $t_{0,\nu}^{\pm}$ be defined as:

$$\begin{aligned}
\|\mathbf{D_u A}(\mathbf{u}_\nu^0)^{-1}\| \leq \quad & \beta_{0,\nu} && \equiv \frac{c}{1-\varepsilon c} < +\infty, && \text{(23a)}\\
\frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|(\mathbf{D_u A}(\mathbf{u}_\nu^0))^{-1}\mathbf{r}_\nu\| \leq \quad & \eta_{0,\nu} && < +\infty, && \text{(23b)}\\
& \gamma_{0,\nu} && \equiv \eta_{0,\nu}\ \beta_{0,\nu}\ \kappa, && \text{(23c)}\\
& t_{0,\nu}^{\pm} && = \frac{1}{\kappa\ \beta_{0,\nu}}\left(1 \pm \sqrt{1-2\gamma_{0,\nu}}\right). && \text{(23d)}
\end{aligned}$$

For each $\nu \geq 1$, we introduce the sequences $\beta_{k,\nu}$, $\eta_{k,\nu}$, $\gamma_{k,\nu}$ and $t_{k,\nu}^{\pm}$ as

$$\begin{aligned}
\gamma_{k,\nu} &= \beta_{k,\nu}\ \eta_{k,\nu}\ \kappa, && \text{(24a)}\\
\beta_{k+1,\nu} &= \frac{\beta_{k,\nu}}{1-\gamma_{k,\nu}}, && \text{(24b)}\\
\eta_{k+1,\nu} &= \frac{\gamma_{k,\nu}\ \eta_{k,\nu}}{2(1-\gamma_{k,\nu})}, && \text{(24c)}\\
t_{k,\nu}^{\pm} &= \frac{1}{\kappa\ \beta_{k,\nu}}\left(1 \pm \sqrt{1-2\gamma_{k,\nu}}\right). && \text{(24d)}
\end{aligned}$$

Taking the limit $k \to \infty$ for the residue Newton-Raphson scheme (13), we have

$$\|\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_\nu)\| \equiv \|\mathbf{u}_\nu^0 - \mathbf{u}_\nu\| \leq t_{0,\nu}^{-} = t_{p_{\nu-1},\nu-1}^{-} \tag{25}$$

and, according to the definition of $\mathbf{u}_\nu^0 \equiv \mathbf{u}_{\nu-1}$:

$$\|\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_{\nu-1})\| \equiv \|\mathbf{u}_{\nu-1}^{p_{\nu-1}} - \mathbf{u}_{\nu-1}\| \leq 2\eta_{p_{\nu-1},\nu-1} \leq \frac{2\eta_{0,\nu-1}}{2^{p_{\nu-1}}}. \tag{26}$$

Kantorovich's Theorem then reads:

Corollary 2.2 *For each index ν, the sequence $(\mathbf{u}_\nu^k)_k$ generated by the scheme (13) converges to the unique solution $\mathbf{u}_\nu \equiv \mathbf{u}(\alpha_\nu)$ of the system:*

$$\mathbf{A}(\mathbf{u}_\nu) - \alpha_\nu \mathbf{r}_{\nu-1} = 0 \quad , \quad \mathbf{r}_{\nu-1} \equiv \mathbf{A}(\mathbf{u}_{\nu-1})$$

in the open ball $B(\mathbf{u}_\nu^0, t_{0,\nu}^+)$ with

$$t_{0,\nu}^+ = \frac{1}{\kappa\beta_{0,\nu}}\left(1 + \sqrt{1 - 2\gamma_{0,\nu}}\right)$$

and where κ, $\beta_{0,\nu}$ and $\gamma_{0,\nu}$ are defined as in (20), (23a) and (23c) respectively.

A sufficient condition for the convergence of (13) can now be stated as it follows.

Proposition 2.3 *A sufficient condition for the sequence $(\mathbf{u}_\nu^k)_{k\geq 0}$ to converge towards $\mathbf{u}(\alpha_\nu)$ is that α_ν satisfies*

$$0 < \frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|\mathbf{r}_\nu\| < \frac{1}{2c}\min\left(\frac{3-\sqrt{5}}{2\kappa c}, t_{0,\nu}^-\right) \equiv \Lambda_\nu, \tag{27}$$

where the constant c has been defined in (16).

Proof. We are seeking for a condition ensuring (21) with $\eta = \dfrac{\varepsilon}{\kappa}$, $\mathbf{u} = \mathbf{u}_\nu^0$ and $\mathbf{u}' = \mathbf{u}(\alpha_\nu)$ that is

$$||\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_\nu)|| < \frac{\varepsilon}{\kappa}. \tag{28}$$

Besides, this must be compatible with (25). In order to achieve (28), we first notice that:

$$\|\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_\nu)\| \leq \|\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_{\nu-1})\| + \|\mathbf{u}(\alpha_{\nu-1}) - \mathbf{u}(\alpha_\nu)\|. \tag{29}$$

Taking into account (26), we may choose $p_{\nu-1}$ so that

$$\frac{2\eta_{0,\nu-1}}{2^{p_{\nu-1}}} < \frac{\varepsilon}{2\kappa}$$

that is,

$$2^{p_{\nu-1}} > \frac{4\eta_{0,\nu-1}\kappa}{\varepsilon},$$

which fixes $p_{\nu-1}$ and also $\mathbf{u}_\nu^0 = \mathbf{u}_{\nu-1}^{p_{\nu-1}}$.
Recall that

$$\forall\alpha \in I_\nu : \quad \mathbf{A}(\mathbf{u}(\alpha)) = \frac{(1-\alpha_\nu)\alpha - (1-\alpha_{\nu-1})\alpha_\nu}{\alpha_{\nu-1} - \alpha_\nu}\mathbf{r}_{\nu-1}, \tag{30}$$

where we set

$$I_\nu \equiv [\min(\alpha_{\nu-1}, \alpha_\nu), \max(\alpha_{\nu-1}, \alpha_\nu)]. \tag{31}$$

Then, (17) yields:

$$\|\mathbf{u}(\alpha_{\nu-1}) - \mathbf{u}(\alpha_\nu)\| =$$

$$\left\| \int_{\alpha_{\nu-1}}^{\alpha_\nu} \frac{d\mathbf{u}}{d\alpha}(\alpha) d\alpha \right\| = \frac{|1-\alpha_\nu|}{|\alpha_{\nu-1}-\alpha_\nu|} \left\| \int_{\alpha_{\nu-1}}^{\alpha_\nu} \mathbf{D_u A}(\mathbf{u}(\alpha))^{-1} \mathbf{r}_{\nu-1} d\alpha \right\|$$

$$\leq \frac{|1-\alpha_\nu|}{|\alpha_{\nu-1}-\alpha_\nu|} \left| \int_{\alpha_\nu}^{\alpha_{\nu-1}} \left\| \mathbf{D_u A}(\mathbf{u}(\alpha))^{-1} \mathbf{r}_{\nu-1} \right\| d\alpha \right| \leq c|1-\alpha_\nu| \|\mathbf{r}_{\nu-1}\|$$

$$= c \frac{|1-\alpha_\nu|}{|\alpha_\nu|} \|\mathbf{r}_\nu\|.$$

Due to (27), we may choose ε such that

$$\frac{2\eta_{0,\nu-1}}{2^{p_{\nu-1}}} = c\frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|\mathbf{r}_\nu\| < \frac{\varepsilon}{2\kappa} < \frac{3-\sqrt{5}}{4\kappa c}. \tag{32}$$

After substitution into (29), this expression can be written as

$$\|\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_\nu)\| \leq 2c\frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|\mathbf{r}_\nu\| < \frac{\varepsilon}{\kappa}, \tag{33}$$

that is (28).
As for (25), the same argument with $t_{0,\nu}^-$ instead of $\dfrac{\varepsilon}{\kappa}$ shows that (27) leads to

$$||\mathbf{u}_\nu^0 - \mathbf{u}(\alpha_\nu)|| \leq 2c\frac{|1-\alpha_\nu|}{|\alpha_\nu|}||\mathbf{r}_\nu|| < t_{0,\nu}^-$$

which yields (25). Moreover, a sufficient condition for the convergence to hold is that

$$0 < \eta_{0,\nu} < \frac{1-\varepsilon c}{2\kappa c}. \tag{34}$$

Indeed: from [10], a sufficient condition for Newton's method to converge at the given step ν is that $\gamma_{0,\nu} < \frac{1}{2}$. Arguing as in (19), we get:

$$\|\mathbf{D_u A}(\mathbf{u}_\nu^0)^{-1}\| \leq \frac{\|\mathbf{D_u A}(\mathbf{u}(\alpha_\nu))^{-1}\|}{1 - \|\mathbf{D_u A}(\mathbf{u}_\nu^0) - \mathbf{D_u A}(\mathbf{u}(\alpha_\nu))\| \|\mathbf{D_u A}(\mathbf{u}(\alpha_\nu))^{-1}\|}$$

$$\leq \frac{c}{1-\varepsilon c} = c' \equiv \beta_{0,\nu} \tag{35}$$

Then, the condition $\gamma_{0,\nu} < \frac{1}{2}$ becomes

$$0 < \eta_{0,\nu} < \frac{1}{2\beta_{0,\nu}\kappa} = \frac{1-\varepsilon c}{2c\kappa}, \tag{36}$$

which is the equation (34). Notice that

$$\eta_{0,\nu} \geq \frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|(\mathbf{D_u A}(\mathbf{u}_\nu^0))^{-1}\mathbf{r}_\nu\|$$

with

$$\|(\mathbf{D_u A}(\mathbf{u}_\nu^0))^{-1}\mathbf{r}_\nu\| \leq c'\|\mathbf{r}_\nu\|.$$

Then, taking into account the equation (35), a sufficient condition reads

$$\frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|\mathbf{r}_\nu\| < \frac{1}{2\kappa}\left(\frac{1-\varepsilon c}{c}\right)^2.$$

Comparing with (27), we get:

$$0 < \frac{|1-\alpha_\nu|}{|\alpha_\nu|}\|\mathbf{r}_\nu\| < \frac{1}{2}\min\left(\frac{\varepsilon}{\kappa c}, \frac{t_{0,\nu}^{-}}{c}, \frac{1}{\kappa}\left(\frac{1-\varepsilon c}{c}\right)^2\right),$$

and

$$\frac{\varepsilon}{\kappa c} < \frac{1}{\kappa}\left(\frac{1-\varepsilon c}{c}\right)^2 \iff \varepsilon^2 - \frac{3\varepsilon}{c} + \frac{1}{c^2} > 0.$$

This is realised as soon as:

$$0 < \varepsilon < \frac{3-\sqrt{5}}{2c},$$

thus finishing the proof. ■

2.2 Pseudo-arclength continuation

For any given $\mathbf{r}$ and assuming that $\operatorname{rank}(\mathbf{D_u A}) = n-1$, we introduce the operator $\mathcal{F} : \mathbb{R}^{n+1} \mapsto \mathbb{R}^n$, as

$$\mathcal{F}(\mathbf{u}(s), \alpha(s); s) \equiv \begin{pmatrix} \mathcal{H}(\mathbf{u}(s), \alpha(s)) \\ \mathcal{N}(\mathbf{u}(s), \alpha(s); s) \end{pmatrix}$$

Following the Implicit Function Theorem, for any given $\mathbf{r}$, the global homotopy (6) can be written as

$$\mathcal{F}(\mathbf{u}(s), \alpha(s); \mathbf{r}) = 0, \tag{37a}$$

$$k_{\mathbf{u}} \cdot \mathbf{u}(s) + k_\mu \mu(\mathbf{u}(s)) = 0. \tag{37b}$$

For some fixed $s_\nu > 0$, consider Newton's scheme (11) written in the equivalent form:

$$\begin{aligned}
&\text{For } \nu = 1, \cdots, N,\\
&\text{For } k = 0, \cdots, p_\nu - 1,\\
&\mathbf{D_u A}_\nu^k(\mathbf{u}_\nu^{k+1} - \mathbf{u}_\nu^k) - (\alpha_\nu^{k+1} - \alpha_\nu^k)\mathbf{r}_{\nu-1} = -\mathbf{A}_\nu^k + \alpha_\nu^k \mathbf{r}_{\nu-1},\\
&\mathbf{D_u N}_\nu^k(\mathbf{u}_\nu^{k+1} - \mathbf{u}_\nu^k) + \mathbf{D}_\alpha \mathbf{N}_\nu^k(\alpha_\nu^{k+1} - \alpha_\nu^k) = -\mathcal{N}_\nu^k,\\
&(\mathbf{u}_\nu^k, \alpha_\nu^k) \leftarrow (\mathbf{u}_\nu^{k+1}, \alpha_\nu^{k+1}).
\end{aligned} \tag{38}$$

For any $1 \le \nu \le N$, the corresponding value of the control parameter μ_ν is such that

$$k_{\mathbf{u}} \,.\, (\mathbf{u}_\nu - \mathbf{u}_{\nu-1}) + k_\mu(\mu_\nu - \mu_{\nu-1}) = 0 \tag{39}$$

where the initialization point $(\mathbf{u}_\nu^0, \alpha_\nu^0) \equiv (\mathbf{u}_{\nu-1}, \alpha_{\nu-1})$ is taken to be solution of (6) with $s \equiv s_{\nu-1}$ and $\mathbf{r} \equiv \mathbf{r}_{\nu-2}$.

In the sequel, we assume that the matrix

$$\mathcal{B}(\mathbf{u}, \alpha) \equiv \begin{pmatrix} \mathbf{D_u A}(\mathbf{u}) & -\mathbf{r} \\ \mathbf{D_u N}(\mathbf{u}, \alpha) & \mathbf{D}_\alpha \mathbf{N}(\mathbf{u}, \alpha) \end{pmatrix}$$

is nonsingular for every $(\mathbf{u}, \alpha) \in \mathbb{R}^n \times \mathbb{R}$ and we assume that there is a constant $c > 0$ such that

$$\|\mathcal{B}^{-1}(\mathbf{u}, \alpha)\| \leq c, \quad \forall (\mathbf{u}, \alpha) \in \mathbb{R}^n \times \mathbb{R}.$$

Since by construction, $\mathcal{F}(\cdot, \cdot; s, \mathbf{r})$ is a homeomorphism $\mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$, for every $s > 0$, (37) admits a unique solution denoted by $(\mathbf{u}(s), \alpha(s))$:

$$\begin{pmatrix} \mathbf{u}(s) \\ \alpha(s) \end{pmatrix} = \mathcal{F}(\cdot, \cdot; s, \mathbf{r})^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

As $\mathcal{F}^{-1}$ is continuously differentiable, $s \mapsto \mathcal{F}(\cdot, \cdot; s, \mathbf{r})^{-1}$ is of class C^1 as well as $s \mapsto (\mathbf{u}(s), \alpha(s))$. In particular, there exists a constant $c > 0$ such that

$$\|\mathcal{B}(\mathbf{u}(s), \alpha(s))^{-1}\| \leq c, \quad \forall s \in \mathbb{R}, \tag{40}$$

and there holds

$$\begin{pmatrix} \mathbf{u}'(s) \\ \alpha'(s) \end{pmatrix} = \mathcal{B}(\mathbf{u}(s), \alpha(s))^{-1} \begin{pmatrix} 0 \\ -\mathbf{D}_s \mathcal{N}(\mathbf{u}(\mathbf{s}), \alpha(s)) \end{pmatrix} \tag{41}$$

Consider the sequence $(\mathbf{y}_\nu^k)_{\nu,k} \equiv (\mathbf{u}_\nu^k, \alpha_\nu^k)_{\nu,k}$ defined for the subdivision of $I \subset \mathbb{R}$:

$$s_0 < s_1 < \cdots < s_N, \quad h_\nu = s_\nu - s_{\nu-1} > 0,$$

by the following scheme. For $\nu = 1, \cdots, N$,

$$\left\{ \begin{array}{rcl} \mathbf{y}_\nu^0 & = & \mathbf{y}_{\nu-1}, \\ \mathbf{y}_\nu^{k+1} & = & \mathbf{y}_\nu^k - \mathcal{B}(\mathbf{y}_\nu^k)^{-1} \begin{pmatrix} \mathbf{A}_\nu^k - \alpha_\nu^k \mathbf{r}_{\nu-1} \\ \mathcal{N}_\nu^k \end{pmatrix}; \ k = 0, 1, \cdots, p_\nu - 1 \\ \mathbf{y}_\nu & = & \mathbf{y}_\nu^{p_\nu}, \quad \mathbf{r}_\nu = \mathbf{A}(\mathbf{u}_\nu). \end{array} \right. \tag{42}$$

Defining the set

$$\mathcal{C} \equiv \{(\mathbf{u}(s), \alpha(s)) \in \mathbb{R}^n \times \mathbb{R}, \quad s \in [s_0, s_N]\},$$

then there exists a compact convex set $D \subset \mathbb{R}^n \times \mathbb{R}$ such that $\mathcal{C} \subset D$. As $\mathbf{D_y}\mathcal{B}$ is continuous, it is also bounded on D and we have

$$\|\mathbf{D_y}\mathcal{B}(\mathbf{y})\| \leq \kappa, \quad \forall \mathbf{y} \equiv (\mathbf{u}, \alpha) \in D,$$

for some constant $\kappa > 0$. As $\mathcal{B}$ is continuous on D, it is also uniformly continuous on D that is:

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall \mathbf{y}, \mathbf{y}' \in D, \quad \|\mathbf{y} - \mathbf{y}'\| < \eta \Longrightarrow \|\mathcal{B}(\mathbf{y}) - \mathcal{B}(\mathbf{y}')\| \le \varepsilon. \tag{43}$$

Furthermore,

$$\|\mathcal{B}(\mathbf{y}) - \mathcal{B}(\mathbf{y}')\| =$$

$$= \left\| \int_0^1 \mathbf{D_y}\mathcal{B}(\mathbf{y} + t(\mathbf{y}' - \mathbf{y}))dt \; (\mathbf{y}' - \mathbf{y}) \right\| \le \kappa \|\mathbf{y} - \mathbf{y}'\|, \quad \forall \mathbf{y}, \mathbf{y}' \in D$$

so that we may choose $\eta = \dfrac{\varepsilon}{\kappa}$. Notice that

$$\begin{aligned} \|\mathcal{B}(\mathbf{y})^{-1}\| &\le \frac{\|\mathcal{B}(\mathbf{y}^0)^{-1}\|}{1 - \|\mathcal{B}(\mathbf{y}^0)^{-1}\| \|\mathcal{B}(\mathbf{y}) - \mathcal{B}(\mathbf{y}^0)\|} \\ &\le \frac{\|\mathcal{B}(\mathbf{y}^0)^{-1}\|}{1 - \varepsilon \|\mathcal{B}(\mathbf{y}^0)^{-1}\|} \le \frac{c}{1 - \varepsilon c} \equiv c' \end{aligned} \tag{44}$$

so that we choose $\varepsilon \in [0, \dfrac{1}{2c}]$.

With each sequence $(\mathbf{y}_\nu^k)_k$, we may associate the quantities $\beta_{k,\nu}$, $\eta_{k,\nu}$, $\gamma_{k,\nu}$ and $t_{k,\nu}^-$ as:

$$\|\mathbf{D_y}\mathcal{B}(\mathbf{y})\| \quad \le \quad \kappa, \quad \forall \mathbf{y} \in D, \tag{45a}$$

$$\|\mathcal{B}(\mathbf{y}_\nu^0)^{-1}\| \le c' \quad \equiv \quad \beta_{0,\nu} < +\infty, \tag{45b}$$

$$\left\| (\mathcal{B}(\mathbf{y}_\nu^0))^{-1} \left(\frac{(1 - \alpha_\nu^0)}{\alpha_\nu^0} \mathbf{r}_\nu, 0 \right)^T \right\| \quad \le \quad \eta_{0,\nu} < +\infty, \tag{45c}$$

$$\gamma_{0,\nu} \equiv \eta_{0,\nu}\beta_{0,\nu}\kappa \le \frac{1}{2}, \tag{45d}$$

$$t_{k,\nu}^\pm = \frac{1}{\kappa \beta_{k,\nu}} \left(1 \pm \sqrt{1 - 2\gamma_{k,\nu}} \right), \tag{45e}$$

$$\gamma_{k,\nu} = \beta_{k,\nu}\eta_{k,\nu}\kappa, \tag{45f}$$

$$\beta_{k+1,\nu} = \frac{\beta_{k,\nu}}{1 - \gamma_{k,\nu}}, \tag{45g}$$

$$\eta_{k+1,\nu} = \frac{\gamma_{k,\nu}\eta_{k,\nu}}{2(1 - \gamma_{k,\nu})}, \tag{45h}$$

where we took into account that

$$\mathcal{N}_\nu^0 = \mathbf{D_u}\mathbf{A}_\nu^0 \delta \mathbf{u}_\nu^0 + \mathbf{D}_\mu \mathbf{A}_\nu^0 \delta \mu_\nu^0$$

$$\equiv \mathbf{D_u}\mathbf{A}_\nu^0(\mathbf{u}_\nu^0 - \mathbf{u}_{\nu-1}^{p_{\nu-1}}) + \mathbf{D}_\alpha \mathbf{A}_\nu^0(\alpha_\nu^0 - \alpha_{\nu-1}^{p_{\nu-1}}) = 0$$

and where

$$\mathbf{y}_\nu^{k+1} = \mathbf{y}_\nu^k - \mathcal{B}(\mathbf{y}_\nu^k)^{-1}\mathcal{F}(\mathbf{y}_\nu^k; s_\nu, \mathbf{r}_{\nu-1}). \tag{46}$$

Now, the same arguments as in the previous section with $\mathbf{D_u A}$, $\mathbf{u}$, α replaced by $\mathcal{B}$, $\mathbf{y}$, s respectively, yield:

Corollary 2.4 *For each index ν, the sequence $(\mathbf{y}_\nu^k)_k$ generated by the scheme (42) converges to the unique solution $\mathbf{y}_\nu \equiv \mathbf{y}(s_\nu) = (\mathbf{u}(s_\nu), \alpha(s_\nu))$ of the system:*

$$\mathbf{A}(\mathbf{u}_\nu) - \alpha_\nu \mathbf{r}_{\nu-1} \quad = \quad 0, \quad \mathcal{K}(\mathbf{u}_\nu) = 0, \quad \mathcal{N}(\mathbf{y}_\nu; s_\nu) = 0, \quad \mathbf{r}_{\nu-1} \equiv \mathbf{A}(\mathbf{u}_{\nu-1})$$

in the open ball $B(\mathbf{y}_\nu^0, t_{0,\nu}^+)$ with

$$t_{0,\nu}^+ = \frac{1}{\kappa\beta_{0,\nu}} \left(1 + \sqrt{1 - 2\gamma_{0,\nu}}\right)$$

and where κ, $\beta_{0,\nu}$ and $\gamma_{0,\nu}$ are defined as in (45a), (45b) and (45d) respectively.

Proposition 2.5 *A sufficient condition for the sequence $(\mathbf{y}_\nu^k)_{k\geq 0}$ to converge towards $\mathbf{y}(s_\nu)$ is that*

$$0 < h_\nu < \frac{1}{2c\|\mathbf{D}_s\mathcal{N}\|} \min\left(\frac{1}{2\kappa}, t_{0,\nu}^-\right) \tag{47a}$$

$$0 < \frac{|1-\alpha_\nu^0|}{|\alpha_\nu^0|} \|\mathbf{r}_\nu\| < \frac{1}{8\kappa c^2} \equiv \Lambda_\nu, \tag{47b}$$

where the constant c has been defined in (40).

Proof. Using the same arguments as used in the proof of the proposition 2.3, the inequality (43) must hold with $\eta = \dfrac{\varepsilon}{\kappa}$, $\mathbf{y} = \mathbf{y}_\nu^0$, $\mathbf{y}' = \mathbf{y}(\alpha_\nu)$, that is:

$$\|\mathbf{y}_\nu^0 - \mathbf{y}(\alpha_\nu)\| < \frac{\varepsilon}{\kappa}.$$

First, notice that:

$$\|\mathbf{y}_\nu^0 - \mathbf{y}(s_\nu)\| \leq \|\mathbf{y}_\nu^0 - \mathbf{y}(s_{\nu-1})\| + \|\mathbf{y}(s_{\nu-1}) - \mathbf{y}(s_\nu)\|. \tag{48}$$

The analogue of (26) holds true, namely:

$$\|\mathbf{y}_\nu^0 - \mathbf{y}(s_{\nu-1})\| \equiv \|\mathbf{y}_{\nu-1}^{p_{\nu-1}} - \mathbf{y}_{\nu-1}\| \leq 2\eta_{p_{\nu-1},\nu-1} \leq \frac{2\eta_{0,\nu-1}}{2^{p_{\nu-1}}}. \tag{49}$$

Therefore, we may choose $p_{\nu-1}$ so that

$$\frac{2\eta_{0,\nu-1}}{2^{p_{\nu-1}}} < \frac{\varepsilon}{2\kappa}$$

that is,

$$2^{p_{\nu-1}} > \frac{4\eta_{0,\nu-1}\kappa}{\varepsilon},$$

which fixes $p_{\nu-1}$ and also $\mathbf{y}_\nu^0 = \mathbf{y}_{\nu-1}^{p_{\nu-1}}$. Moreover, (41) yields:

$$\|\mathbf{y}(s_{\nu-1}) - \mathbf{y}(s_\nu)\| = \left\|\int_{s_{\nu-1}}^{s_\nu} \mathbf{y}'(s)ds\right\| = \left\|\int_{s_{\nu-1}}^{s_\nu} \mathcal{B}(\mathbf{y}(s))^{-1} \left(0, -\mathbf{D}_s\mathcal{N}(\mathbf{y}(\mathbf{s}))\right)^T ds\right\|$$

$$\leq \int_{s_{\nu-1}}^{s_\nu} \left\|\mathcal{B}(\mathbf{y}(s))^{-1}\right\| \|\mathbf{D}_s\mathcal{N}\| ds \leq c\|\mathbf{D}_s\mathcal{N}\| h_\nu. \tag{50}$$

Then, (47a) implies that we may choose ε so that

$$\frac{2\eta_{0,\nu-1}}{2^{p_{\nu-1}}} = c\|\mathbf{D}_s\mathcal{N}\|h_\nu < \frac{\varepsilon}{2\kappa} < \frac{1}{4\kappa c},$$

in accordance with (22).
After substitution in (48) and taking into account of (49) and (50) it comes:

$$\|\mathbf{y}_\nu^0 - \mathbf{y}(s_\nu)\| \leq 2c\|\mathbf{D}_s\mathcal{N}\|h_\nu < \frac{\varepsilon}{\kappa}, \tag{51}$$

which is (43).
The analogue of (25) reads

$$\|\mathbf{y}_\nu^0 - \mathbf{y}(s_\nu)\| \equiv \|\mathbf{y}_\nu^0 - \mathbf{y}_\nu\| \leq t_{0,\nu}^- = t_{p_{\nu-1},\nu-1}^-. \tag{52}$$

Then, the same argument with $t_{0,\nu}^-$ instead of $\dfrac{\varepsilon}{\kappa}$ yields

$$0 < h_\nu < \frac{1}{2c\|\mathbf{D}_s\mathcal{N}\|}\min\left(\frac{\varepsilon}{\kappa}, t_{0,\nu}^-\right).$$

Moreover, arguing as in the previous section we find that the requirement (34) remains true, that is

$$0 < \eta_{0,\nu} < \frac{1-\varepsilon c}{2\kappa c}. \tag{53}$$

Recall that

$$\eta_{0,\nu} \geq \left\|(\mathcal{B}(\mathbf{y}_\nu^0))^{-1}\left(\frac{(1-\alpha_\nu^0)}{\alpha_\nu^0}\mathbf{r}_\nu, 0\right)^T\right\|$$

so that $\eta_{0,\nu}$ may be chosen as:

$$\left\|(\mathcal{B}(\mathbf{y}_\nu^0))^{-1}\left(\frac{(1-\alpha_\nu^0)}{\alpha_\nu^0}\mathbf{r}_\nu, 0\right)^T\right\| \leq \frac{c}{(1-\varepsilon c)}\left(\frac{|1-\alpha_\nu^0|}{|\alpha_\nu^0|}\|\mathbf{r}_\nu\|\right) \equiv \eta_{0,\nu}.$$

This implies due to (47b):

$$\eta_{0,\nu} \leq \frac{1}{8\kappa c(1-\varepsilon c)}$$

while (22) yields $1-\varepsilon c > \frac{1}{2}$ and thus

$$\frac{\eta_{0,\nu}}{1-\varepsilon c} < \frac{1}{8\kappa c(1-\varepsilon c)^2} < \frac{1}{2\kappa c}$$

which is (53). ■

References

[1] Allgower, E.L., Georg, K. (1990). *Numerical continuation methods.* Springer series in Computational Mathematics Vol. 13, Springer-Verlag.

[2] E. Doedel, E., Keller, H., Kern'evez, J., (1991). J. Numerical analysis and control of bifurcation problems. I: Bifurcation in finite dimensions, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 1: 493-520.

[3] E. Doedel, E., Keller, H., Kern'evez, J. (1991). Numerical analysis and control of bifurcation problems. II: Bifurcation in infinite dimensions, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 1: 745-772.

[4] Guckenheimer, J., Worfolk, P. (1993). Dynamical systems: some computational problems. In: *Bifurcations and Periodic Orbits of Vector Fields*, Montreal, 1992. Schlomiuk, D., ed. Vol. 408 of NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., Dordrecht: Kluwer, pp. 241-278.

[5] Garcia, C.B., Zangwill, W.I. (1981). *Pathways to Solutions Fixed Points and Equilibria*. Englewood Cliffs, NJ: Prentice-Hall.

[6] Kantorovich, L., Akilov, G. (1964). *Functional Analysis in Normed Spaces*, Oxford: Pergamon Press.

[7] Keller, H. B. (1977). Numerical solution of bifurcation and nonlinear eigenvalue problems, In: Rabinowitz, P.H., ed. *Applications of Bifurcation Theory*. New-York: Academic Press, pp. 359–384.

[8] Keller, H. B. (1978). Global homotopies and Newton methods. In: deBoor, C., Gollub, G. H. eds. *Recent Advances in Numerical Analysis*. New-York: Academic Press, pp 73–94.

[9] Keller, H. B. (1987). *Lectures on Numerical Methods in Bifurcation Problems*. Berlin, Heidelberg, New-York: Springer-Verlag.

[10] Ortega, J., Rheinboldt, W. (1970). *Iterative Solution of Nonlinear Equations in Several Variables*, New-York: Academic Press.

[11] Rheinboldt, W.C. (1986). *Numerical Analysis of Parametrized Non-Linear Equations*. New-York: J. Wiley.